



Stephan Czysch

So optimieren Sie mit dem Site Explorer der Bing-Webmaster-Tools kostenlos Ihre Website

Bing treibt die Weiterentwicklung seiner Webmaster-Tools weiter voran. Seit der umfangreichen Vorstellung der Toolsammlung in Ausgabe 63 hat sich nochmals etwas getan: Der bereits aus der „alten“ Version bekannte Site Explorer wurde umgezogen. Im Zusammenspiel mit der URL-Überprüfung erhalten Webmaster einen Einblick in die von Bing gecrawlten und indextierten Adressen einer Website. Erfahren Sie in diesem Artikel von Stephan Czysch, was das neue alte Tool kann.

DER AUTOR



Stephans Leidenschaft gilt besseren Websites. Seit seinem Abschied aus dem von ihm gegründeten Online-Marketing-Team von Dept (ehemals Trust Agents) unterstützt er Unternehmen und Inhouse-Teams auf Projektbasis. Stephan spricht regelmäßig auf Konferenzen zu Themen wie Online-Marketing-Strategien und datengetriebenes SEO. Sein Wissen können Sie in Form seiner Bücher konsumieren – oder besuchen Sie eines seiner Seminare.

Die Themen Crawling und Indexierung beschäftigen viele Webmaster. Besonders Google legte lange Zeit einen Mantel des Schweigens darüber, welche Adressen ihren Weg in den Google-Index fanden. Zwar gab Google die Anzahl der indextierten Seiten in den damaligen Google-Webmaster-Tools an, konkrete Infos zu einzelnen Adressen fehlten allerdings. Anders handhabte Bing dieses Thema: Die Suchmaschine aus dem Haus Microsoft gewährte mit dem **Site Explorer** bereits seit einigen Jahren Einblicke in den Index.

Zum Zeitpunkt der Vorstellung der Bing-Webmaster-Tools Beta in Ausgabe 63 war es noch nicht so weit, doch seit Mitte Oktober steht der **Site Explorer in neuer Optik und mit**

neuen Funktionen zur Verfügung. Zeit also, sich mit diesem Tool auseinanderzusetzen. Und wenn wir schon bei Neuerungen sind: Seit Anfang Oktober heißt der Suchdienst nicht mehr Bing, sondern Microsoft Bing und bekam zudem ein neues Logo spendiert. Es passiert einiges im Microsoft-Suchuniversum!

Vorneweg: Suchmaschinen crawlen nicht wie dein Crawling-Tool

Wer denkt, dass Suchmaschinen sich ähnlich wie Crawling-Tools, beispielsweise Screaming Frog oder Ryte, von einem Startpunkt aus mittels Links durch eine Website crawlen, der täuscht sich. Natürlich sind Verlinkungen, und

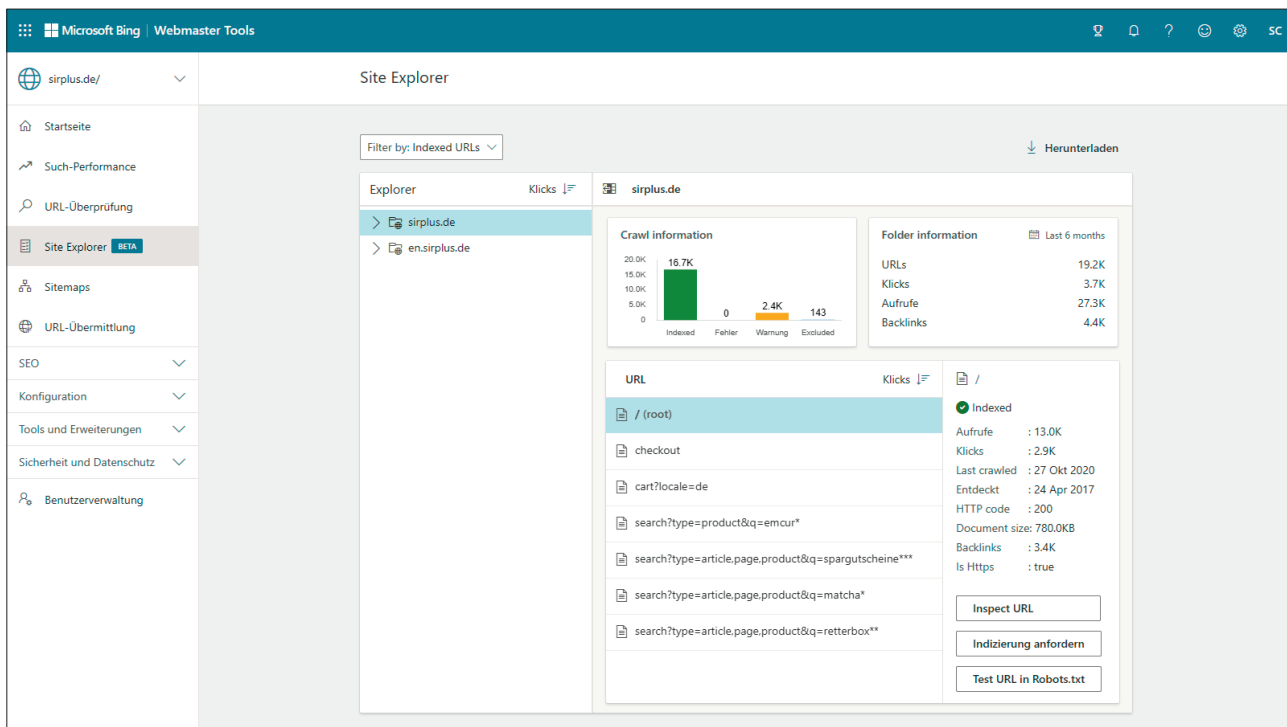


Abb. 1: Auf der obersten Ebene stehen im Site Explorer die unterschiedlichen von Bing gefundenen Subdomains einer Website

dazu zählen z. B. auch Nennungen von Adressen in einer XML-Sitemap, für das Auffinden von Adressen sowie das Crawling wichtig, doch Suchmaschinen crawlen eben nicht alle von einer gerade analysierten Adresse verlinkten Seiten direkt im Anschluss. Ein solches Vorgehen wäre viel zu ineffizient, da sich ein Großteil der Webseiten zwischen Abrufen durch Suchmaschinen überhaupt nicht verändert.

Um die eigenen Crawling-Kapazitäten möglichst effizient einzusetzen, berechnen Suchmaschinen wie Google und Bing Crawling-Prioritäten pro Adresse und rufen Adressen desselben Webauftritts unterschiedlich häufig ab. Indikatoren, die für eine (temporär) häufigere Crawlfrequenz sorgen können, sind beispielsweise:

- » Die stärkere Verlinkung einer Seite
- » Eine hohe Veränderungshäufigkeit in der Vergangenheit
- » Ein verändertes lastmod-Datum in einer XML-Sitemap

Schauen wir uns zwei konkrete Beispiele an, die zeigen, dass unterschiedliche Crawling-Frequenzen sinnvoll sind. Bei einer Tageszeitung ändert sich der Inhalt der Startseite

deutlich häufiger als bei einer privaten Website. Entsprechend wird die Startseite der Tageszeitung tendenziell häufiger von Suchmaschinen-Robots angesteuert. Allerdings gilt dies wahrscheinlich weniger für die Datenschutzbestimmung der Zeitung. Folglich wird diese nur sehr unregelmäßig gecrawlt, obwohl sie vermutlich im Footer und damit stark verlinkt ist. Zudem kommt dadurch mit jeder neuen URL ein weiterer Link hinzu.

Auf welche Adressen greifen Suchmaschinen also zu und welche Seiten sind überhaupt über Suchmaschinen zu finden? Denn das Crawling ist nur der erste Schritt, dem sich, wenn eine Seite über Suchmaschinen gefunden werden soll, zwingenderweise die Indexierung anschließen muss. Einblick in dieses Thema rund um die Bing-Suche bietet der Site Explorer, der das **Pendant zum Bericht Abdeckung der Google Search Console** darstellt.

Da sich das Crawling zwischen Suchmaschinen nicht gravierend unterscheidet, sind die in den jeweiligen Webmaster-Tools zu findenden Daten übertragbar. Hier lohnt sich z. B. ein Blick auf die Struktur der gecrawlten

Adressen. Lassen sich Muster erkennen? Erzeugt z. B. ein bestimmter Seitenbereich 404-Fehler? Solche Informationen helfen Ihnen, schnell(er) Licht ins Dunkel rund um das Crawling und Crawling-Fehler zu bringen.

Kurz und bündig: Websites in Bing-Webmaster-Tools verifizieren

Vermutlich haben Ihnen die Artikel rund um Bing Maps und Bing-Webmaster-Tools in den letzten Website-Boosting-Ausgaben Lust auf die Suchmaschine von Microsoft gemacht. Deshalb haben Sie Ihre Websites wohl bereits in den Bing Webmaster-Tools bestätigt. Falls dies nicht der Fall ist, hier die kurze Anleitung zur Bestätigung Ihrer Website:

- » Loggen Sie sich mit Ihrem Google-Search-Console-Konto bei Bing ein – alle in der GSC bestätigten Websites sind dadurch auch bei Bing bestätigt.
- » Legen Sie ein Konto bei Bing an und integrieren Sie den Verifizierungsschlüssel der Bing-Webmaster-Tools auf Ihre Website.

Der Site Explorer der Bing-Webmaster-Tools

Im Site Explorer dreht sich alles um **(gefundene) und meistens bereits gecrawlte Adressen**, deren **Indexierung** und eventuell vorhandene Probleme. Die Bandbreite reicht von weiterleitenden Adressen über URLs, die Serverfehler zurückliefern. Den Begriff „Explorer“ verwendet Microsoft vermutlich deshalb, weil Sie sich wie beim Windows-Betriebssystem durch die Struktur der Website klicken können.

Die Daten werden im Tool **pro Subdomain zusammengefasst**. Als Webmaster profitieren Sie davon, dass sich Bestätigungen bei Bing immer auf die gesamte Website beziehen. Sie sehen also die Daten jeglicher Subdomains, ohne diese separat bestätigen zu müssen.

Google hat dies mit den Domain-Properties erst seit Februar 2019 im Angebot. Davor war die heute „URL-Properties“ genannte Bestätigung der Standard. Diese hatte (und hat) allerdings den Nachteil, dass die Daten nur für Adressen unterhalb der bestätigten Struktur zur Verfügung stehen. Haben Sie Ihre Website unter <https://www.bestätigt>, dann sind Daten unter <https://blog.bestätigt> nicht enthalten. Diese Logik hat allerdings bei einigen Anwendungsfällen Vorteile und mit der Kombination aus URL- und Domain-Properties haben Sie bei Google Search Console den besten Überblick über Ihre Website.

Besonders für SEO-Agenturen liefert der Site Explorer eine schnelle Übersicht darüber, welche Subdomains bei einem neuen Kunden zum Einsatz kommen. Damit eine Subdomain im Bericht auftaucht, muss in diesem Fall Bing die Subdomain natürlich über Verweise kennengelernt haben.

Innerhalb der Google Search Console müssten Sie, um unterschiedliche Hostnamen zu finden, im Abdeckungs-Bericht über die oberhalb der Tabelle platzierten Filter einzelne URL-Muster

ausschließen. Hier kann allerdings die Beschränkung auf maximal 1.000 Zeilen schnell einen Strich durch die Rechnung machen.

Werfen Sie einen Blick auf Abbildung 1, um sich den Aufbau des Site Explorers anzuschauen. Das Tool ist in folgende Bereiche aufgeteilt:

- » Links: Der **Explorer-Bereich** zur Navigation innerhalb von Subdomains und Verzeichnissen. Ein Pfeil zeigt an, ob sich innerhalb der gewählten Struktur weitere auswählbare Verzeichnisse befinden.
- » Rechts oben: **Statistiken und Informationen** der gewählten Subdomain oder Verzeichnis-Struktur.
- » Rechts unten: Der **URL-Bereich** mit den einzelnen Adressen sowie deren individuelle Daten samt drei Analyse-Links. Diese führen zur URL-Prüfung, zur Neu-Indexierung oder zum robots.txt-Tester.

Dazu stehen oberhalb des Arbeitsbereichs die **Filterfunktion** sowie der **Download** zur Verfügung. Zusätzlich können die Daten sowohl im Explorer als auch im URL-Bereich **sortiert** werden. Standardmäßig sind die Adressen nach den erzielten Klicks in den letzten sechs Monaten sortiert. Alternativ lassen sich die Adressen von A-Z oder nach Aufrufen, also erzielten Impressionen in den Suchergebnissen, sortieren. Zur Erinnerung: Eine Impression ist die Anzeige einer Adresse Ihres Webauftritts auf einer von Nutzern aufgerufenen Suchergebnisseite.

Bevor es in die Details geht, schauen wir auf die Statistiken. Diese werden von Bing in Crawl-Informationen (links) sowie Ordner-Informationen (rechts) unterteilt.

Bei den **Crawl-Informationen** geht es um die Anzahl an Adressen, die in vier verschiedene Segmente runtergebrochen werden:

- » Indexierte Seiten
- » Adressen mit Fehlern
- » Adressen mit Warnungen

- » Ausgeschlossene URLs

Leider führt ein Klick auf diese einzelnen Segmente nicht dazu, dass die Daten im URL-Bereich entsprechend eingeschränkt werden. Dazu kann, mit Abstrichen, die Filterfunktion verwendet werden.

Direkt neben den Crawl-Informationen sind die **Statistiken** zu sehen.

Auch diese sind in vier Datenpunkte aufgeteilt:

- » Die Anzahl der Adressen im Verzeichnis
- » Die Anzahl der erzielten Klicks in der Bing-Suche
- » Die Anzahl der Impressionen, also wie häufig die Adressen zu Suchanfragen erschienen sind (im Deutschen ungenau mit „Aufrufe“ übersetzt)
- » Die Anzahl der Backlinks auf die Adressen im Verzeichnis

Oberhalb der Tabelle ist die bereits angesprochene **Filterfunktion** zu finden. Hier stehen folgende Filterungen zur Auswahl:

- » Alle Adressen (Im Tool: „All URLs“)
- » Indexierte Seiten (Im Tool: „Indexed URLs“) – das ist der Standardwert
- » Nicht mehr funktionierende URLs mit HTTP-Statuscodes 404 und 410 (Im Tool: „URLs with Crawling Issues (dead link – http 404-410“)
- » Nicht mehr funktionierende URLs mit HTTP-Statuscodes 403 und 5xx (Serverfehler) (Im Tool: „URLs with Crawling Issues (server errors – http 403, 5xx“)
- » Per robots.txt ausgeschlossene Adressen (Im Tool: „URLs robots.txt disallowed“)
- » Über Noindex ausgeschlossene Adressen (Im Tool: „URLs with Noindex tag“)
- » Weiterleitende Adressen (Im Tool: „URLs redirecting“)
- » Nicht-kanonische Adressen (Im Tool: „URLs Canonical source“)
- » URLs mit Verstößen gegen die Bing-Webmaster-Guidelines (Im Tool:

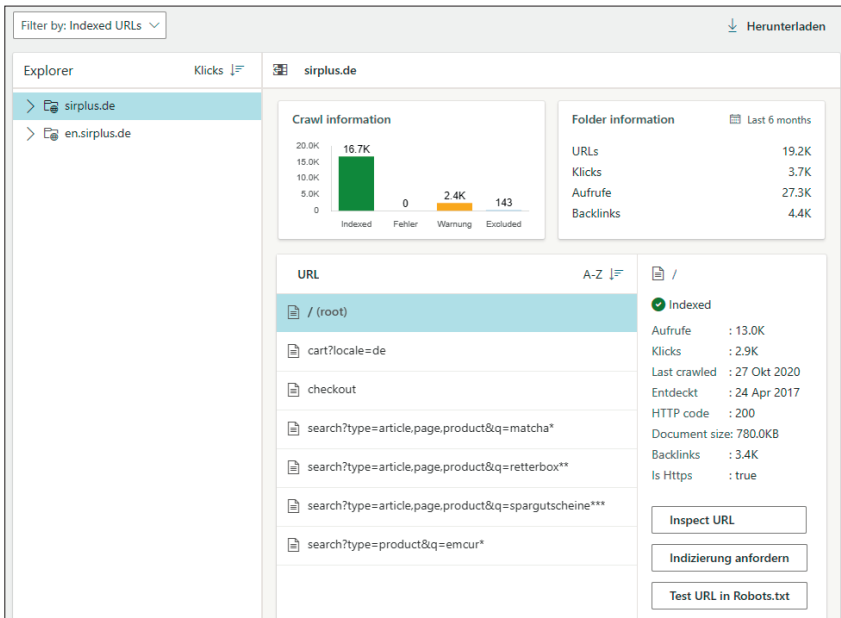


Abb. 2: Zu jeder Adresse werden im unteren rechten Bereich die von Bing gespeicherten Infos angezeigt

klar: Sie können Adressen identifizieren, die (nicht) erfolgreich verarbeitet werden konnten, und diese anschließend überarbeiten.

Mit dem Site Explorer arbeiten

Standardmäßig ist als Filter oberhalb der Daten die Option „Indexierte URLs“ ausgewählt. Klicken Sie sich mit diesem Filter durch den im Tool links platzierten Explorer-Bereich, dann sehen Sie **einen Auszug der zur gewählten URL-Struktur gehörenden und von Bing indixierten Seiten.**

- » „URLs with guideline issues“
 - » Bisher nicht gecrawlte Adressen (Im Tool: „URLs selected not yet crawled“)
 - » Adressen mit Schadsoftware (Im Tool: „URLs with malware“)
 - » Adressen mit anderen Problemen (Im Tool: „URLs with other issues“)
- Einige dieser Gruppen gibt es in

identischer Form in der Index-Abdeckung Google Search Console, andere sind Bing exklusiv wie die Auflistung von Adressen, die gegen die Webmaster-Richtlinien verstoßen. Die Bing-Webmaster-Richtlinien können Sie unter <http://einfach.st/binggl> nachlesen. Anhand der Filter werden die Einsatzmöglichkeiten des Site Explorers

Zwar begrenzt Bing im Gegensatz zu Google nicht die Anzahl der Zeilen auf 1.000, doch auch über Bing sind nicht alle **indexierten Seiten** zu erhalten. Dies gilt sowohl für die im Interface angezeigten Adressen als auch für den Download. Denn dieser extrahiert nur die gerade angezeigten URLs samt den dazugehörigen Daten.

Wer darauf hofft, dass die Auswahl der höchsten Ebene alle URLs einer Subdomain zurückliefert, der täuscht

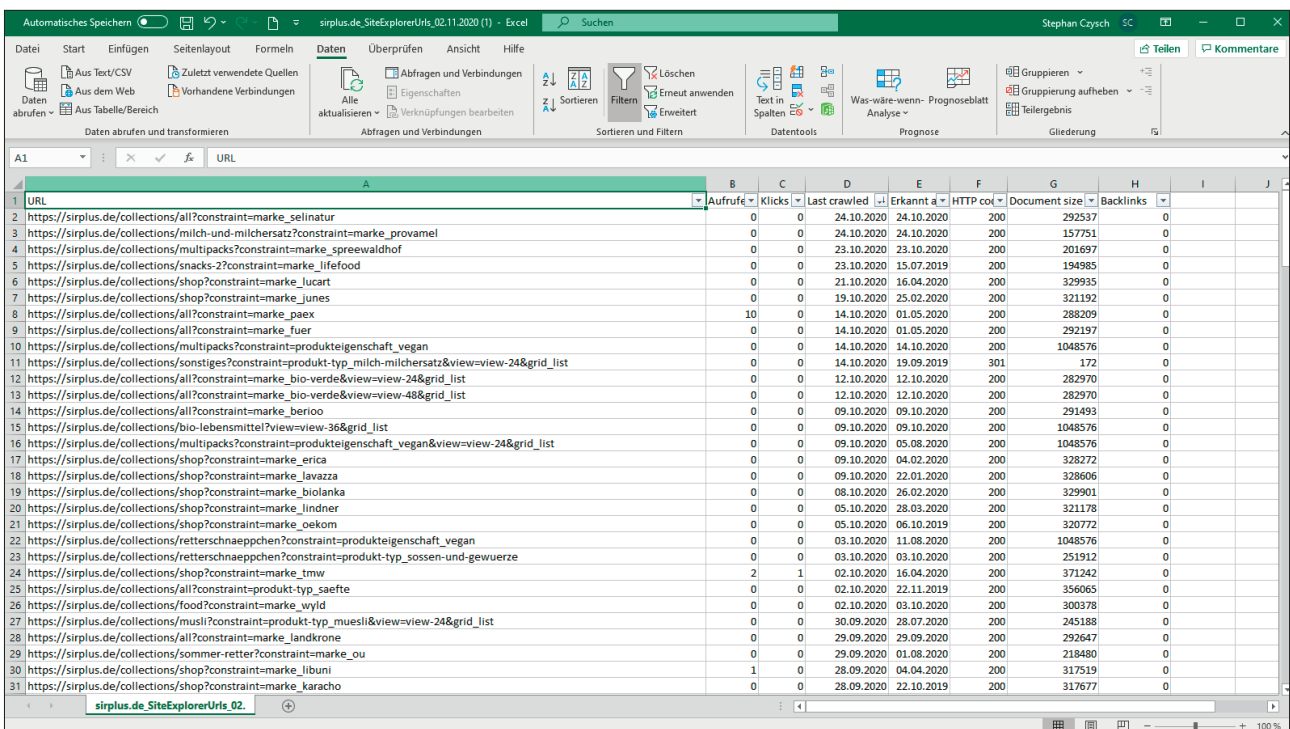


Abb. 3: Der Export enthält bis auf die HTTPS-Info alle Daten der im Interface gerade angezeigten Seiten

sich. Denn jegliche URLs in (Unter-) Verzeichnissen werden erst dann angezeigt, wenn das entsprechende Verzeichnis angeklickt wurde. So sind unter Abbildung 2 nur die sieben indexierten Seiten unterhalb der Subdomain <https://sirplus.de/> zu sehen, da diese eben nicht in einem Verzeichnis liegen.

Um mehr Infos zu einer einzelnen Seite zu erhalten, muss diese innerhalb des URL-Bereichs ausgewählt werden. Als Information liefert Ihnen Bing die folgenden Daten zurück:

- » Wurde die Seite erfolgreich aufgerufen und ist sie aktuell indexiert?
- » Wie häufig ist sie in den Bing-Ergebnissen in den letzten sechs Monaten erschienen („Aufrufe“)?
- » Wie häufig wurde sie angeklickt?
- » Wann wurde die Website zuletzt gecrawlt?
- » Wann hat Bing die URL erstmals gesehen?
- » Welcher HTTP-Statuscode wurde beim letzten Aufruf zurückgeliefert?
- » Wie groß ist die Seite?
- » Wie viele eingehende Links zeigen auf die URL?
- » Läuft die Seite unter dem HTTPS-Protokoll?

Bis auf die Information zum HTTPS-Protokoll sind all diese Informationen im Download enthalten. Diese fehlende Information können Sie allerdings mit einem Blick auf die URL erhalten, da Bing die volle Adresse anzeigt und sowohl HTTP- als auch HTTPS-Adressen auflistet.

Wofür können Sie den Site Explorer nutzen?

Wie Sie nun wissen, dreht sich im Site Explorer alles um die Bing bekannten Adressen eines Webauftritts. Konnte die URL fehlerfrei abgerufen werden? Ist sie indexiert? Verstößt sie gegen die Richtlinien? Ist die Adresse nicht kanonisch? Wie groß ist das Dokument? Wie viele Adressen leiten

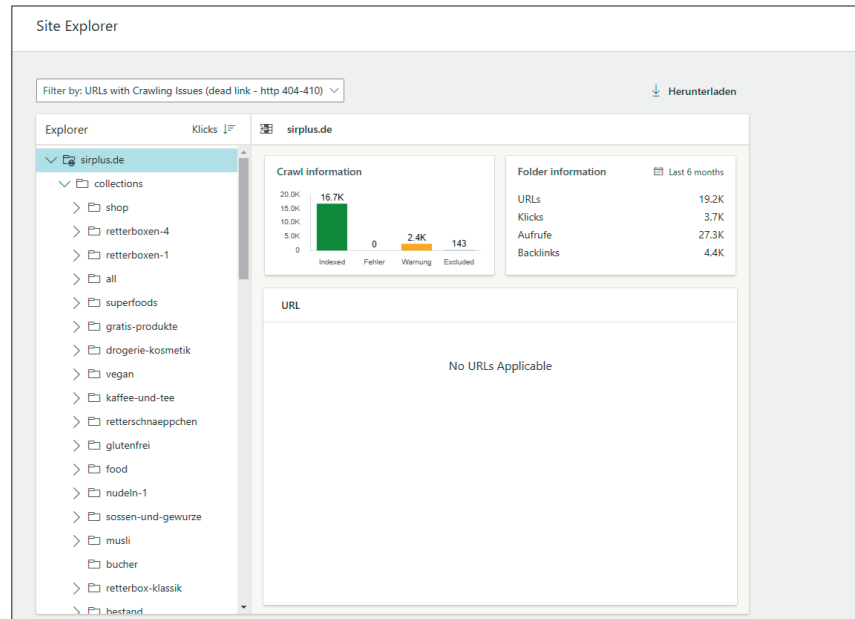


Abb. 4: Leider zeigt Ihnen Bing bei Auswahl der Subdomain nicht alle vorhandenen fehlerhaften Adressen an

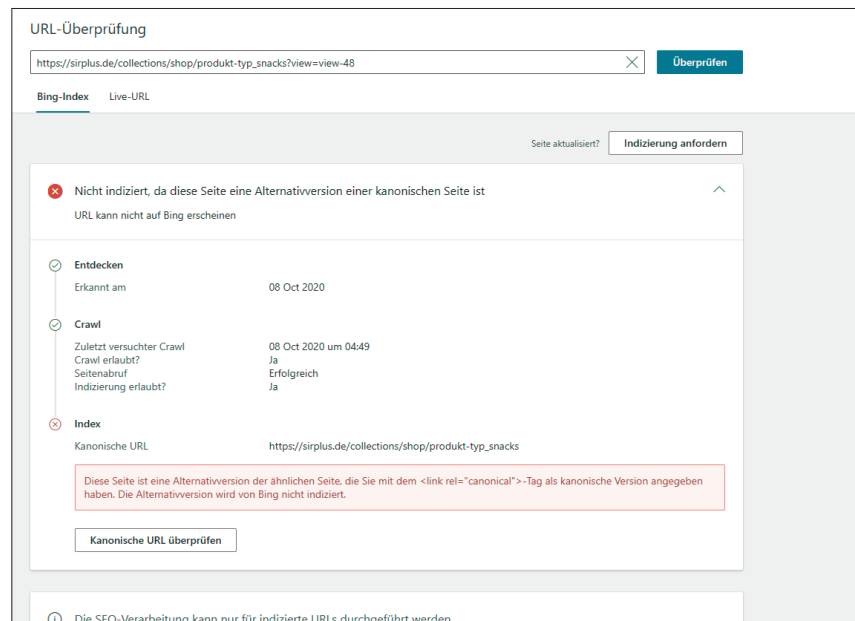


Abb. 5: Die in der URL-Überprüfung analysierte Adresse wurde von Bing als „nicht-kanonisch“ eingestuft

auf andere Seiten weiter? Welche Seiten erzeugen einen Crawling-Fehler und müssen ggf. weitergeleitet werden? Die Liste der vom Site Explorer gelieferten Antworten ist lang und die Aufzählung nur ein erster Überblick. Durch die vordefinierten Filter können Sie die Daten segmentieren und sich das anzeigen lassen, was Sie gerade interessiert.

Die Explorer-Logik stößt hier (leider) von Zeit zu Zeit an Grenzen. Denn wenn Sie alle z. B. 404- oder 410-URLs sehen möchten, dann reicht eine Auswahl der Subdomain nicht aus. Wie Sie in Abbildung 4 sehen können, entsprechen auf der ausgewählten Ebene keine Adressen diesem Problem. Doch in den Unterverzeichnissen des Webauftritts sind einige Fehler zu finden.

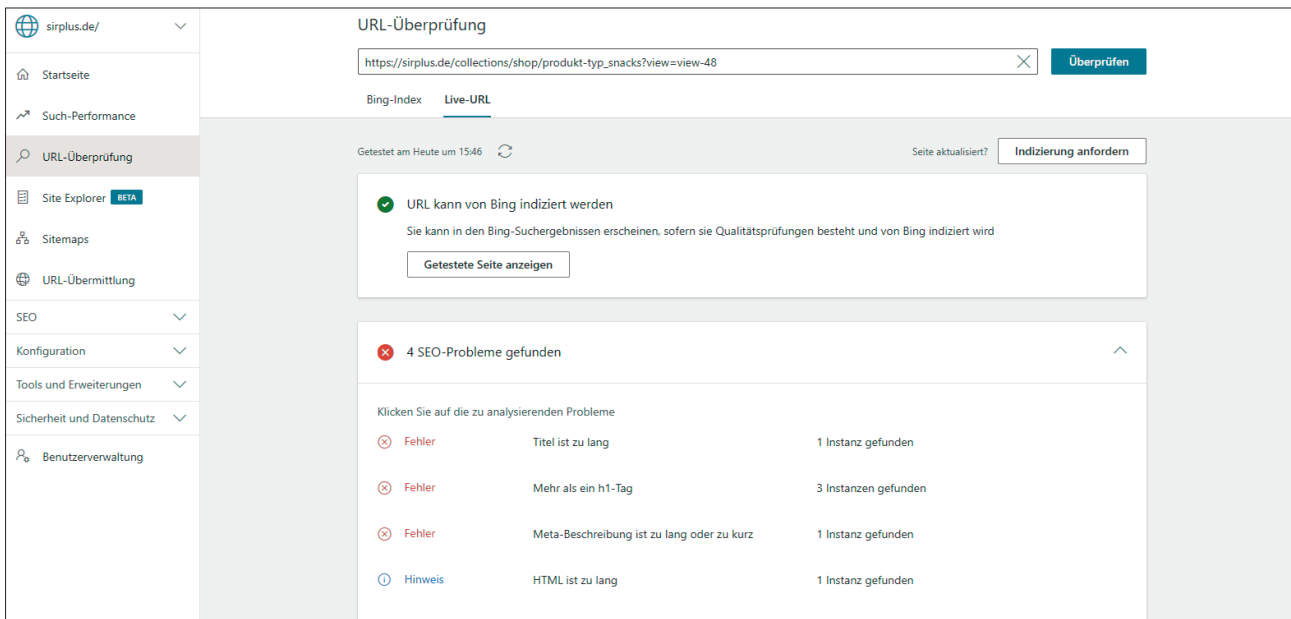


Abb. 6: Mit dem Live-Test crawlt Bing eine Seite neu und identifiziert dabei direkt SEO-Probleme sowie strukturierte Daten; das Tool kann zudem für eine Neu-Indexierung verwendet werden

Auf mich wirkt das wenig durchdacht und ich persönlich würde mich freuen, **wenn ein Verzeichnis alle darunterliegenden Daten zurückliefern würde**. Vielleicht passt Microsoft das noch an.

Auf den ersten Blick können zudem die Statistiken oberhalb des URL-Bereichs verwirren. Denn diese Daten passen sich nicht dem gewählten Filter an. Aus Abbildung 4 lässt sich also nicht ablesen, wie viele Klicks auf (aktuell von Bing so eingestufte) Fehlerseiten gingen. Die Statistiken beziehen sich stets auf den gesamten Ordner – unabhängig vom Status der Adressen. Eine Möglichkeit der Sortierung von URLs nach Status dürfte die Usability zudem nochmals verbessern. Denn ansonsten müssen Sie immer mit dem Filter arbeiten – und dieser spült Sie bei Auswahl immer zurück auf die oberste Ebene des Berichts.

Wenn Sie mehr über eine einzelne Adresse erfahren möchten, dann können Sie diese über den Link „Inspect URL“ mit der **URL-Überprüfung** analysieren.

Die URL-Überprüfung zur Fehlersuche nutzen

Als weitere Neuerung kann die URL-Überprüfung begrüßt werden. In diesem Tool werden einzelne Adressen analysiert. Neben dem Abruf der beim letzten Crawl von Bing festgestellten Daten kann eine **Live-Analyse** durchgeführt werden. Analog zur URL-Überprüfung von Google werden dabei direkte Verbesserungsvorschläge zurückgeliefert.

In Abbildung 5 ist das Testergebnis einer von Bing im Site Explorer als fehlerhaft markierten Adresse zu finden: Die Adresse wurde zum Zeitpunkt des letzten Zugriffs von Bing zwar erfolgreich abgerufen, allerdings ist die Seite nicht-kanonisch. Die Detailinfos werden in der URL-Überprüfung angezeigt. Da die in Abbildung 5 angezeigte Adresse nicht indexiert wurde, zeigt Bing keine Verbesserungsvorschläge basierend auf den Bing-Indexdaten an.

Die Verbesserungsvorschläge sind aktuell auf **SEO-Best-Practices** sowie **strukturierte Daten** beschränkt. Bei erfolgreich abrufbaren Seiten (HTTP-Statuscode 200) werden diese Tests beim Klick auf „Live-URLs“ ad hoc

durchlaufen. Und siehe da: Die überprüfte Adresse aus Abbildung 5 weist im Live-Test Probleme auf. Ein Klick auf einen der Fehler oder Hinweise öffnet dabei ein neues Fenster mit dem von Bing abgerufenen Quelltext. Dort lässt sich der Fehler oder Hinweis direkt über die Suchfunktion identifizieren.

Die SEO-Potenziale sind alternativ über den Einzeltest „Website-Prüfung“ im Navigationsbereich SEO zu finden. Zudem werden in „SEO-Berichte“ verschiedene Adressen in die von Bing ausgewerteten Fehlerkategorien sortiert.

In Kombination können die beiden Tools Ihnen dabei helfen, Crawling-Fehler zu identifizieren und die Indexierung zu überwachen sowie zu optimieren. Gibt es Adressen, die Sie nicht für Suchmaschinen erreichbar machen wollen? Sind wichtige Seiten versehentlich ausgeschlossen? Der Site Explorer hilft Ihnen dabei, Ihre Website zu verbessern. Werfen Sie deshalb von Zeit zu Zeit einen Blick auf die Daten und die Bing-Webmaster-Tools. ¶